View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository@Nottingham

Reputation Aware Obfuscation for Mobile Opportunistic Networks

Milena Radenkovic, Abderrahim Benslimane, Derek McAuley

Abstract: Current anonymity techniques for mobile opportunistic networks typically use obfuscation algorithms to hide node's identity behind other nodes. These algorithms are not well suited to sparse and disconnection prone networks with large number of malicious nodes and new opportunistic, adaptive. So, new, opportunistic, adaptive fully localized mechanisms are needed for improving user anonymity. This paper proposes reputation aware localized adaptive obfuscation for mobile opportunistic networks that comprises of two complementary techniques: opportunistic collaborative testing of nodes' obfuscation behaviour (OCOT) and multidimensional adaptive anonymisation (AA). OCOT-AA is driven by both explicit and implicit reputation building, complex graph connectivity analytics and obfuscation history analyses. We show that OCOT-AA is very efficient in terms of achieving high levels of node identity obfuscation and managing low delays for answering queries between sources and destinations while enabling fast detection and avoidance of malicious nodes typically within the fraction of time within the experiment duration. We perform extensive experiments to compare OCOT-AA with several other competitive and benchmark protocols and show that it outperforms them across a range of metrics over a one month real-life GPS trace. To demonstrate our proposal more clearly, we propose new metrics that include best effort biggest length and diversity of the obfuscation paths, the actual percentage of truly anonymised sources' IDs at the destinations and communication quality of service between source and destination.

Index Terms—Mobile communication systems, Privacy

1. Introduction

Rapid expansion of free social networking applications and mass scale distribution of user generated content is often contributed to pervasive mobile devices that allow mobile users to generate and consume digital content. Underlying networking protocols [20], [14], [17], [33], [10], [16], [19], [30], [28], [27] increasingly provide support for large scale content sharing that utilize both opportunistic networks among mobile devices and connections to cellular network and access points when within their range.

There is a wide range of opportunistic content multimedia sharing systems that have been discussed in the literature such as podcasting [13] or interest based content sharing application for public transportation [15]. For example, bulky nature of multimedia content typically generated by mobile devices when certain events happen or around interesting geographical locations, can benefit from opportunistic networking as useful extension of the core infrastructure because that can carry large amount of traffic on its own [28]. Additionally, these networks have been used to create new applications, such as epidemic voting [10] and social media [10], [12]. PeopleNet is architecture for information search in a distributed manner [8]. In [9] the authors propose a socio-aware overlay for an opportunistic pub/sub communication service. An interest based city-wide opportunistic content dissemination system has been studied in [16]. Various architectures for email delivery from Internet to and within an opportunistic networking domain were discussed in [17]. In [19] a scalable content dissemination system with dynamic content is discussed that exploits both infrastructure and opportunistic contacts.

Despite offering advantages for mobile users in terms of increased coverage and lower cost, the proposals that include opportunistic forwarding of user data have not yet been widely adopted. One of the key reasons for that is that users are inherently reluctant to interact with strangers or third parties, due to privacy concerns of social networking [20]. Work in [18], [22], [31], [34], [6], [21] focuses on how new approaches on improving privacy of users of opportunistic networks by obfuscating their identity, location or adding incentives for good behavior. Newly emerging work [20] proposes a new architecture that addresses the issues of lack of trust, delivery latency, loss of user control, and user privacy by combining the advantages of decentralized storage and opportunistic communications by bridging the gap between the user and the mobile social networking friends who are not in the vicinity of the user. However, none of this work considers adaptive, self-organised privacy where trust levels of nodes can change depending on their behaviour which is the work we do in this paper.

In this paper we explore how to design adaptive

- *Milena Radenkovic is with the School of Computer Science, The University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK, E-mail: milena.radenkovic@nottingham.ac.uk*
- *Abderrahim Benslimane is with the Computer Science Laboratory, University of Avignon, 339 Chemin des Meinajaries BP 91228, Avignon 84911, France, E-mail: abderrahim.benslimane@univ-avignon.fr*
- *Derek McAuley is with the School of Computer Science, The University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK, E-mail: Derek.mcauley@nottingham.ac.uk*

xxxx-xxxx/0x/$xx.00 © 200x IEEE Published by the IEEE Computer Society

reputation aware anonymous overlays in mobile opportunistic networks that manage complex trade-offs between quality of anonymisation of sending nodes and the quality of service the nodes receive (e.g. answers to their queries without being deanonymised). Managing this tradeoff is important as the better the obfuscation of the source ID is, the more difficult it is to identify it and route to it; and thus the success ratio of answered queries decreases and the delays between source and destination increase. We assume that both the number and the distribution of malicious nodes that can reveal source identity can vary, good nodes can at times become malicious and vice versa, the underlying network can get very sparse and the user interests may also change. Because of this, it is important that the anonymisation overlay is responsive to the underlying topology, users' interests and varying number of malicious nodes due to their behavioral changes in order to significantly improve nodes' anonymity and privacy.

We propose to incorporate novel Opportunistic Collaborative Obfuscation Technique (OCOT) with a flexible multidimensional K-anonymity obfuscation technique AdaptAnon (AA) [22] in order to identify and build adaptive anonymisation overlays (paths) to hide the source's identity from malicious nodes. OCOT-AA manages to discover malicious and trusted nodes with high accuracy and efficiently while providing high level of communication quality of service between sources and destinations.

The rest of the paper is organised as follows. Related work is given in the supplementary file. Section 2 describes our threat model and provides high-level overview of privacy and forwarding in our opportunistic network proposal. Section 3 gives overview of a set of heuristics that are at the core of our OCOT-AA proposal and describes our algorithm in detail. Section 4 provides overview of our network model, real-life trace, defines multiple evaluation metrics, describes experiment scenario including benchmark and comparative protocols, and then discusses results from extensive evaluation of OCOT-AA. Section 5 gives conclusions.

The contributions of this paper are multifold and include:

- Identification of new privacy threats implicit in utilizing services in mobile opportunistic networks.
- Proposal for opportunistic collaborative obfuscation testing (OCOT) technique that allows intelligent, efficient and accurate fully distributed reputation building based on the nodes' obfuscation behaviour. We define efficiency as the percentage of testing time versus connectivity time between the nodes; it is denoted as "testing cost". We define accuracy in terms of how fast the detection of the majority of malicious nodes and trusted nodes is.
- Proposal for adaptive anonymisation technique that dynamically combines three types of implicit fully localized heuristics rooted in social complex graph theory that can track anonymisation history of all the nodes, adaptively select suitable anonymisers while taking into account that nodes may provide false information on their own heuristics. We define success of anonymisation in three following ways: percentage of source IDs obfuscated in the destination, maximum best effort achieved length of the obfuscation path (K), and maximum best effort diversity of the nodes in it (D).
- Demonstration of practical feasibility of our proposal in terms of identification and mitigation of attacks over real-life trace highly challenged real life trace while preserving high level of communication quality of service.

Our privacy analytics is useful for large-scale and online environments – e.g. in our real time obfuscation work [22] we have used live real time GPS traces of San Francisco cabs [2] to show real-time visualisation of anonymisation parameters for AdaptAnon in ONE simulator [25]. We chose this trace as it is currently one of the largest real-world traces available (involving over 540 nodes), spanning over multiple months with consistently highest frequency of movement updates. We envisage that our work can be easily integrated into fast evolving pervasive systems that require self-organising security capabilities driven by predictive patterns of complex dynamic graph theory.

2. Privacy and Forwarding in Mobile Opportunistic Networks

2.1. Threat Model

There is a range of source obfuscation protocols that have been proposed for opportunistic networks with no infrastructure [6], [18], [21], [31]. They differ in how they determine optimal obfuscator by relying on a range of heuristics based on interest or connectivity similarities. Although these metrics are reasonable in a friendly wireless environment, they are not accurate in a hostile scenario and may lead to selecting malicious nodes as forwarders.

The most trivial attack is that malicious node waits passively until other nodes select them as forwarding hops and then obtain their identities. But, to increase the probability of being selected as next hops for forwarding and therefore to improve their chances of revealing the sources' IDs, malicious nodes can advertise false information about their own social network parameters such as centrality, similarity, betweenness, tie-strength, obfuscation history etc. This indirect attack aims to attract the traffic and queries forwarded to them and allows them to identify large number of sources' identities.

Testing the actual obfuscation behaviour of potential obfuscator nodes and assigning obfuscation reputation to them before selecting them is important and reduces the chances of malicious nodes being selected as obfuscators. However there is a range of attacks that can happen while these tests are being carried out. We assume that even the nodes that have been tested and reached a reputation value high enough to be considered "trusted" to perform obfuscation can become malicious at any point of time by starting to reveal other nodes' identities. In this paper we will introduce techniques that will opportunistically keep

testing the reliability of obfuscation behaviour of all the nodes (those considered to be "trusted", malicious and tested). We will discuss a range of specific attacks during and after testing and how OCOT handles them in 3.2.

In this work we will not consider network level attacks such as black or grey hole attacks as the attacking nodes target to reveal the identity of the source nodes and not to disrupt their network traffic.

2.2. Reputation Aware Adaptive Obfuscation

This section proposes OCOT-AA that combines two complementary techniques: adaptive obfuscation forwarding technique (AdaptAnon (AA) [21], [22]) and novel Opportunistic Collaborative Obfuscation Testing (OCOT) technique. OCOT-AA aims to fully locally test obfuscation behaviour of neighbouring nodes, gather and provide evidence of how good reputation in terms of obfuscation behaviour a node has **each time before** it is selected to be on the obfuscation path. The tested node can "pass" the tests and become "trusted" only if it gets tested by a certain number of "friend or "trusted" nodes and if the result of the aggregated test values exceeds "Threshold" (e.g. by using Eigen Vector Reputation Centrality). We describe this in more detail in Section 3.

We refer to *opportunistic* as the ability of the nodes to test whenever and wherever they can. This is important in the face of potentially sparse and dynamic topologies. We refer to *collaborative* as to the requirement that the nodes need to collaborate with other nodes in order to perform the tests of obfuscation behaviour. Collaborative and opportunistic paradigms are typically assumed to be opposites of each other. To the best our knowledge this is the first proposal where we combine these two approaches in the reputation building schemes (depicted in Figure 1).

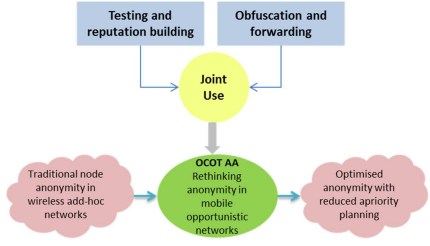

Figure 1. Reputation aware adaptive obfuscation in mobile opportunistic networks

OCOT-AA exploits dynamic label matching, temporal analysis of contact patterns and intelligent obfuscation in order to "hide" a source node among not only more nodes and but also more diverse nodes which provides stronger anonymity [5]. Section 3.1 introduces in detail our new proposal for OCOT technique while section 3.2. discusses its robustness and 3.3 gives its pseudo code. In 3.4 we describe novel flexible, multi-dimensional approach to K-anonymity that enables reputation aware opportunistic identification and selection of the overlay anonymisation nodes in order to allow for better trade-off management between the maximum best effort length of the obfuscation path and the maximum best effort diversity of the nodes on it while avoiding malicious nodes and not degrading communication success ratio and delays.

Figure 1 shows schematic representation of reputation aware adaptive obfuscation for mobile opportunistic networks (OCOT-AA). OCOT-AA enables a paradigm shift from traditional node anonymity in wireless ad-hoc networks to optimised anonymity with reduced *apriori* planning via the joint use of testing and reputation building and obfuscation and forwarding.

Depending on *where* the reputation value is stored, *when* it is exchanged, and between *which* nodes, there is a range of possible solutions with different pros and cons. If we consider two radically different scenarios: first, tested node stores all reputation values assigned to it by the testing nodes and shares them when it encounters other nodes; and second, each testing node keeps its view of reputation values it assigns to the nodes it tested and exchanges them with other testing nodes to resolve the final reputation. In the first case, as the tested node stores complete history of its encounters and test results, the reputation building process always converges and the network overhead is low. In the second case, each node has its local view of the encountered nodes' obfuscation behaviours based on its own experience and shares it with the other testing nodes. In this case, different parts of the network may resolve reputation for the same node differently, reputation convergence may be very slow and the network overheads might be very high as all nodes need to aggregate their knowledge by exchanging their information on encountering each other. This fully decentralised reputation building is not suitable for sparse hostile mobile opportunistic networks [35]. We therefore choose to design and build on the assumption of the first scenario where each node stores its reputation data about its own obfuscation behaviour that is signed by the nodes that tested it. We expand on this in Section 3.

3. Reputation Aware Obfuscation Testing and Adaptive Anonymisation for Mobile Opportunistic Networks

3.1 Opportunistic Collaborative Obfuscation Testing

We propose a self-organized reputation management technique coupled with certificate exchange that improves node identity obfuscation in mobile opportunistic networks. Our proposal assumes a set of F "friends" certified by the initial CA (e.g. any online social network such as Facebook). These nodes then act as second level CAs themselves and certify the public keys of the "trusted" nodes with their own certificates. We refer to the nodes that gained their reputation by performing reliable obfuscation tasks during testing as "trusted" nodes. Both "friends" and "trusted" nodes opportunistically test all other nodes by requesting from them to perform obfuscation of dummy messages where one of the "trusted" nodes is the source and the other one is the destination. These tests serve to identify whether

nodes perform obfuscation and hide the real ID of the sending node or if they reveal it to the destination. Note that there must be at least two "testing" nodes collocated with the "tested" node in order for testing to be possible. When the destination receives a message from the tested node, it sends the ID that the tested node sent to it, back to the source. The functional overview of the OCOT testing scheme is shown in Figure 2.

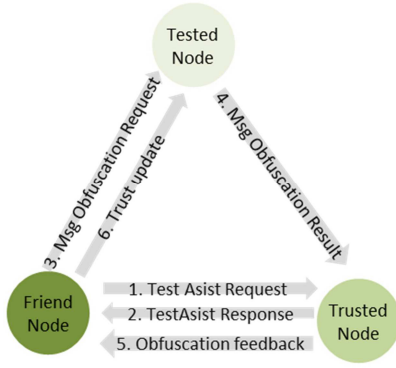

Figure 2. Testing scheme (OCOT)

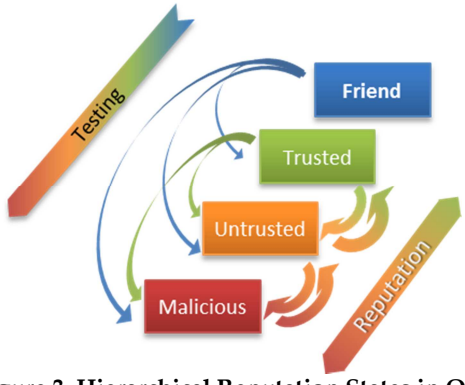

Figure 3. Hierarchical Reputation States in OCOT

Figure 3 shows hierarchical reputation state changes of nodes: when the node's reputation increases it moves towards the "trusted" state and when it decreases: it moves towards "malicious" node state. Any unknown node starts with reputation value of 0.1, the increase coefficient is 1 and the decrease coefficient is 4 where the coefficients are weighted by the trust value of the testing node. The reputation value for a "non-friend" node cannot go above 0.95 or go below 0. Figure 3 also shows that "friend and "trusted" nodes can perform testing on "trusted" and "malicious" nodes but not vice versa.

During the time of testing the tested node does not know if the request for obfuscation is genuine and contains real data or an obfuscation test with dummy data. This is because the sending node may be considered to be the last node on the obfuscation path that aims to deliver the message to the destination with obfuscation path length equal to 1. Even if, after the testing has been completed, the tested node understands that it has been tested (e.g. by statistically analysing the obfuscation requests it received), it cannot change the opinions of the testing nodes. The test results are signed by the testing nodes' private keys and cannot be modified by anyone. On the other hand, if the tested node refuses to participate in the test by either forwarding the data to another destination node different from the destination requested by the sending node, or by refusing to forward the data altogether, the testing nodes will not rank it highly and it will not be able to build its reputation value to become a "trusted" node (therefore no node will use it for the real forwarding and obfuscation requests). It is possible that a tested node refuses to store the result of test when it failed to satisfy the obfuscation request. Even though this can bring a short term benefit to the tested node, but the node's reputation will quickly expire below the trusted threshold due to the testing timeout. Note that a node can preserve its "trusted" status only by continuously responding correctly to anonymisation requests.

3.2 Robustness of Testing in OCOT

One particularly interesting question is how much damage a node that has passed all the tests and became "trusted" can do when it decides to stop obfuscating the source nodes, and what mechanisms OCOT provides to minimise the negative impacts of this. OCOT can revoke a "trusted" node's reputation in two ways: first the signed keys of the "trusted" nodes expire after a relatively short period of time, and second "friend" and "trusted" nodes keep performing obfuscation tests on the "trusted" nodes. If a "trusted" node fails a test, the testing node will issue a new certificate to the tested "trusted" node in which the testing node will overwrite its reputation value in their certificate to much lower (or zero) reputation values. Note, we do not revoke a node's certificate itself but only modify its reputation value. Consequently, even though an "untrusted" node will be able to store its test results to another node, its opinion will not be taken into account.

In order to account for attacks launched by the "trusted" nodes during the testing phase we aggregate reputation values of all involved testing nodes. For example, when a "friend" node (with reputation 1) initiates testing with a "trusted" node (with reputation 0.8), "trusted" node may send false information to the "friend" node, the "friend" node signs the tested node with the obfuscation results weighted by 0.8 (1*0.8).

Another critical question refers to how OCOT can prevent the corrupt "trusted" node from revealing the reputation results, node IDs and timestamps to the services and to the "tested" nodes. To counteract this, in OCOT, the "friend" nodes' certificate subject cannot be easily mapped to the "friend" node real ID because the "friend's pseudo ID used in the tests is different from its real ID. We assume that all "friend" and "trusted" nodes have short term pseudo IDs that can change often (i.e. a node can provide a different pseudo ID for each new test) in addition to its real ID. This means that the compromised "trusted" node may be able to determine that node X1 had tested it at time t1 but it will not know who X1 node is.

When there are multiple nodes to be tested in the neighbourhood, OCOT follows either a target based policy, such as the node that is the closest to passing the reputation threshold gets tested first, or age based policy -

the one that has been tested the longest time ago gets tested first. In highly challenged environments where the expected connectivity durations are short and the density of nodes is high, it is important to be efficient and test nodes with higher priority first as their involvement in the obfuscation will result in significant difference to the performance of OCOT.

3.3. Pseudo code of OCOT and mathematical notation

The pseudo code of the OCOT algorithm is shown in Figure 4. It begins with identifying the nodes in the neighbourhood that could perform the testing and those that should be tested (lines 2, 3): The pseudo-code then moves to how the testing nodes are chosen and paired up to test tested nodes. This is important in order to determine if testing is feasible at a given time and to prevent repetition of sequential coupling of the same testing/tested nodes (lines 5-7, 11-13). We also ensure that we do not test the same node far too often (lines 8-10).When the three nodes that participate in a test are allocated, second part of the code outlines the testing process. It shows how the source of the testing node generates a pseudo ID (line 14), and uses it to initiate a request for obfuscation testing to the selected "friend" or "trusted" node (line 15). It then initiates a request for obfuscation to the tested node asking it to forward its pseudo ID to the "trusted"/"friend" node. It waits for the tested node to perform the anonymisation and for the "friend"/"trusted" node to forward the result of the obfuscation back to it (line 16). After it receives the response it generates a value for the Reputation status of the tested node based on its opinion of the tested node performance, signs it digitally and stores the updated Reputation value onto the tested node (lines 17-20).

ALGORITHM TestNeighbours

```
1.  FUNCTION TestNeighbours(List Contacts) {
2.      List TestingNodes = extractTestingNodes(Contacts);
3.      List TestedNodes = extractTestedNodes(Contacts);

4.      IF (TestingNodes.Length() AND TestedNodes.Length()) THEN
5.          FOR (int c=0; c < TestedNodes.Length(); c++) DO
6.              Node TestedNode = TestedNodes(c);
7.              Node TestingNode = TestingNodes(c%TestingNodes.Length());
8.              IF (TimeNow – TestedNode.TimeLastTest) < TestInterval THEN
9.                  NEXT
10.             END IF
11.             IF TestedNode == TestingNode  THEN
12.                 TestingNode = TestingNodes((c+1)%TesingNodes.Length());
13.             END IF
14.             FakeNodeId = generateTempRandomNodeId();
15.             requestTestingCooperation(TestingNode, FakeNodeId);
16.             requestObfuscation(TestedNode, TestingNode, FakeNodeId);
17.             float TestResult = recvTestResult(TestingNode);
18.             float ReputationValue =
                        updateCommulativeReputation(TestedNode, TestResult, TestingNode. reputation_value);
19.             X509Sign(ReputationValue);
20.             storeReputation(TestedNode, ReputationValue);
21.         END FOR
22.     END IF
23. END FUNCTION
```

Figure 4. OCOT Algorithm

This is more formally annotated in Formula 1.

$$T_{A,B}(U) = \begin{cases} A, B & if\ \exists\ A,B \\ T''_{A,B}(U) + T'_{A,B}(U) & if\ P_{T'} \geq I_T \\ 1 & \end{cases} \quad (1)$$

where $T_{A,B}(U)$ is the present reputation rating of node U by nodes A and B, $T''_{A,B}(U)$ is the reputation rating recorded during the previous test. The conditions for the present testing are expressed as $if\ \exists\ A, B$ standing for the availability of "trusted" testing nodes A and B, and $if\ P_{T'} \geq I_T$, representing the time period since the previous testing $P_{T'}$ as longer or equal to the pre-configured testing period I_T. The length of the testing period can play significant role in the performance of our proposed approach. We discuss this in detail and present graphs in the supplementary file.

3.4 Obfuscation and Forwarding

As the accuracy of the test results is more important for obfuscation testing than the time factor, we use the average values of weighted reputation results without discounting histories. Thus we resolve reputation as the weighted sum of all opinions divided by the total number of nodes as

$$Tm = \frac{\sum_{i=1}^{|V|} w_i \cdot T_i}{|V|} \qquad |v \geq 8 \quad (2)$$

where T_i is opinion of the node i for the reputation value of node m, and w_i is the weight of this opinion that depends on the reputation value that node i has at the time of signing node m's reputation value, and v is the

total number of nodes that tested node m. The minimum required number of testing nodes, 8, has been experimentally determined to achieve best trade-off between quality of anonymisation and quality of service. When the reputation values are updated and before the source forwards its request to the service, the source initiates obfuscation process that aims to hide its identity by enabling opportunistic identification and selection of the trusted anonymisation overlay. This overlay manages complex trade-offs between reliability, length of the obfuscation path and the diversity of the nodes on it while not degrading success ratio and delays. Prioritisation of overlay nodes is based first on their reputation values, and then on connectivity patterns, and interest profile similarity, and anonymisation history. More detailed discussion on our obfuscation heuristics is given in the supplementary file but here we give a brief overview of three main classes of heuristics**:** 1) locally evaluated heuristics driven by the network topology and contact history analysis include nodes' "betweenness" centrality [1], [11], social "similarity" [1], [11], and tie strength relationships [1], [11], between the nodes; 2) heuristics driven by interest and user profile analysis include the degree of profiles' similarity nodes share when they meets based on the number of matched profile attributes versus the number of total attributes (L), denoted as *LabelSimilarity*.

$$LabelSimilarity = \frac{|L(N) \cap L(M)|}{|L(N) \cup L(M)|} \quad (3)$$

3) Combining social connectivity driven and profile driven metrics allows OCOT-AA to increase the length of the anonymisation path (K) compared to using only one of these two metrics alone because it allows more options for the selection of the next hop anonymisation node. However this can still result in predictable choices of the nodes in the anonymisation layer and thus have negative impact on the anonymisation quality. In order to counterbalance the decidability of the previous two heuristics, we propose the third type of heuristics that is driven by the analytics of the anonymisation history performed by every node and for every potential anonymising node. This allows OCOT-AA to increase the diversity of the nodes in the K overlay in order to improve the utilisation of the overlay nodes. Each node calculates the ratio of the number of times the next hop has been used by the given origin and by all other sources in order to be able to make less greedy decisions.

$$\text{NodeRatio (N)} = \frac{\text{NodeAnonRqst(N)}}{\sum_{i \in N} \text{NodeAnonRqst(i)}} \quad (4)$$

Each node also keeps the ratio of the number of times the next hop has been on the anonymisation path for the particular Service and has been used for all other services.

$$\text{ServiceRatio (S)} = \frac{\text{ServiceAnonRqst(S)}}{\sum_{j \in S} \text{ServiceAnonRqst(j)}} \quad (5)$$

Finally, each node keeps track of the ratio of the number of times the next hop has been used to anonymise this source for this service, and the number of times it has anonymised other nodes for this service.

$$ServiceNodeRatio\ (N,S) = \frac{|NodeAnonRqst(N) \cap ServiceAnonRqst(S)|}{|ServiceAnonRqst(S)|} \quad (6)$$

Note that in this paper we assume equal weights between heuristics but it is also possible to use different weighing models in order to prioritise some criteria over the others if that is suitable. As we assume unfriendly environments with varying percentage of malicious nodes, OCOT constantly re-evaluates the nodes trust values and ranks them so it results in most accurate and least predictable chosen nodes. Each node scans the neighbourhood and performs testing as described in Figure 3 and ranks the nodes according to their reputation values. Only if a node is a "trusted" node, i.e. its reputation value exceeds the predefined reputation threshold it may be selected as a next hop for obfuscation. After it has been determined that a node is a "trusted" node, forwarding is described in more detail and pseudo code is listed in the supplementary file.

4. Evaluation Methodology and Results

4.1. Network Scenario

Mobile opportunistic networks enable mobile users to participate in various social interactions with applications such as content distribution and micro-blogs. Mobile opportunistic networks allow direct one-hop communications between devices carried by people or vehicles while on the move. In this way the nodes are involved in participatory interactions with their surrounding without having to rely on wireless infrastructure. Because of their open and distributed nature, securing user interactions relies rather on trust than hard cryptography [4], [19], [32], [33]. Trust in this context is typically based on past users' interactions and behaviour such as in reputation systems relying on ratings that may be based on "belonging" to online social networks and on other users opinions. For example social online networks are utilized where a set of nodes coming from the same online social network is assumed to explicitly trust each other. By leveraging the social online networks for secure pairing of online nodes with wireless nodes, we propose to use social trust bootstrapping that cannot be compromised together with implicit reputation management that relies on nodes obfuscation behaviour, connectivity and interest analytics. As the nodes in mobile opportunistic networks can use only the nodes they encounter for data transmission the connectivity analysis (rather than mobility) is key to how much data nodes can communicate with the others. [33] defines connectivity duration, isolation duration and average number of neighbours as key metrics that influence communication feasibility and quality in these networks.

Previous research has performed comparisons of various real-life traces and identified substantial differences in connection duration, isolation duration, and number of neighbours. San Francisco Cab Trace [2] has been shown to be the most challenging trace with very short connectivity durations [1], very long isolation time [1] and low number of neighbours [1] compared to other social traces (such as Sassy[7] and Infocom[14]). For

example, SF Cab [2] traces exhibit predominantly short contact durations (a mean of 31sec) while Infocom 2006 displays substantially longer contact durations (a mean of 3min). In this work we choose to focus on reporting on SF traces as most challenging dataset currently available.

4.2. Real life data set description and analysis

We performed extensive evaluation of OCOT-AA in ONE[25] using live GPS traces of 540 San Francisco cabs, logged approximately every 10 seconds, over a period of 30 days. We downloaded the most recent at the time of writing traces for the period of September 20th 2012 to October 20th 2012 via the Cabspoting.org API. These traces are part of the Cabspotting project [2] that aim to infer and visualise Taxi Cab collocation information from GPS coordinates in the San Francisco Bay Area. We have assumed that two cabs are collocated if their physical distance is less than 50 meters and used a time interval of 60 seconds. This trace has shown to exhibit long periods of disconnections, short periods of connectivity and islands of connectivity that are rarely populated by more than two nodes.

4.3. Evaluation metrics/criteria

We perform evaluation across ten metrics described below.

We first analyse the efficiency of OCOT in order to better understand what costs in terms of traffic and delays it imposes to the obfuscation and forwarding. It is very important that testing does not take negative impact on the data traffic during the connectivity times by taking large proportion of connectivity time and resources for testing. Similarly, given potentially highly sparse topology, it is crucial that the malicious nodes are discovered quickly. Our experiments show both high efficiency in terms of testing resources and high speed of discovering the malicious and trusted nodes. This is due to our reputation converging very fast and being accurate. We use five different metrics: number of tests performed, trusted and malicious detection rate, testing cost, detection rate and number of different nodes performing testing, achieved K and D versus sparseness of topology to show effectiveness and accuracy of our proposal. These metrics are discussed in detail in the supporting file. In this paper we define five core metrics for evaluating effectiveness of OCOT-AA that includes: success ratio and delays of answered queries, length and diversity of the anonymisation path (K and D) and failure rate. We define D as

$$1 - \frac{Kp_t \cap Kp_{\Delta t}}{Kp_t \cup Kp_{\Delta t}} \qquad (7)$$

where Kp_t is the set of nodes on the K path at the time of a query and $Kp_{\Delta t}$ is the set of the node the K path at the time of the next query. We define anonymisation failure as the real identity of the source getting revealed to the destination. We define failure rate as

$$\frac{M_f}{M_t} \qquad (8)$$

where M_f is the number of messages with failed obfuscation and M_t is the total number of messages delivered to destination.

4.4. Experiment set up and comparison with benchmark and state of the art protocols

During experiments, when we analyse the influence of changing numbers and distribution of malicious nodes, we keep percentage of "friend" nodes to 5% while we increase percentage of malicious nodes with the following steps of increase: 5%, 10%, 20%, 30% and 40%. We argue that this choice of scenarios allows us to test our protocol in a range of different scenarios which is important as they can be applicable to the real world scenarios. We perform comparative analysis of OCOT-AA and two other competitive protocols SRR [19], where trust is assumed between nodes belonging to same social communities, and SimBetTS [11], where trust is based the similarity of social connectivity between nodes. To allow fair comparison, we extend SRR and SimBetTS to support obfuscation in the nodes that are elected to be on an anonymisation path by the source. We show that OCOT-AA manages highly dynamic trade-offs between directionality and diversity, as well as anonymity and quality of service. We show that explicit testing is essential for high levels of anonymity and assumptions of implicit trust conveyed in utilising online social networks or similarity of opportunistic contacts based analysis can be damaging as it results in high levels of failed anonymisation.

4.5 Results

4.5.1 Success Ratio and Delay

Figure 5 shows the success ratio over increased number of queries. AdaptAnon [21] has the highest success ratio, followed closely by SimBetTS and OCOT-AA 5% M, OCOT-AA 10% M OCOT-AA 20% M. OCOT-AA 30% M OCOT-AA 40% M and SRR [19] manage less than half the success ratio compared to the first three.

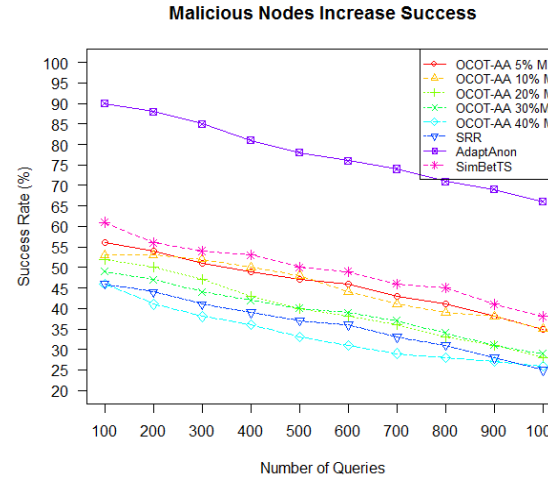

Figure 5. Success ratio vs. increasing number of queries

AdaptAnon performs the best because it utilises two (social and encounter-based) mechanisms to determine the next anonymising node on its path. SimbetTS is second because it efficiently exploits the high degree of

encounter similarity over SF taxi trace that has social complex graph properties. OCOT-AA 5% M follows as its testing method allows it to monitor obfuscation behaviour and detect malicious nodes quickly and find paths that are similar to AdaptAnon's and Simbets's. We observe that OCOT-AA 10% M, OCOT-AA 20% M, OCOT-AA 30% M, OCOT-AA 40% M achieve success ratios that are inversely proportional to the percentages of malicious nodes. SRR manages the worst success ratios as it cannot detect and utilise nodes that have matching social roles in a sparse topology.

Figure 6 shows delays (in seconds) for varying percentages of malicious nodes over increasing number of queries. The three worst performing protocols with highest delays are OCOT-AA 40% M, OCOT-AA 30% M, and OCOT-AA 20% M that manage delays from 2000 to 2800 seconds. This is due the large number of malicious nodes that cannot be utilized for testing and obfuscation. The best performing protocol is SimbetTS because it targets the most direct route to the destination. OCOT-AA 5% M, SRR, OCOT-AA 10% M and AdaptAnon have very similar performance between themselves and are only about 400 seconds above SimBetTS for low sending rates and about 800 seconds above for high sending rates. They manage performance about 800 seconds below OCOT-AA 40% M, OCOT-AA 30% M, and OCOT-AA 20% M for low sending rates and are nearly the same as them for high sending rates. OCOT adds only less than 5 min compared to the anonymisation protocols without testing, has similar delay to epidemic benchmark routing protocol without any anonymisation and is better than random walk This shows that delays between sources and destinations are mainly due to the sparse and disconnected underlying topology rather than the testing and intelligent anonymisation that OCOT-AA employs.

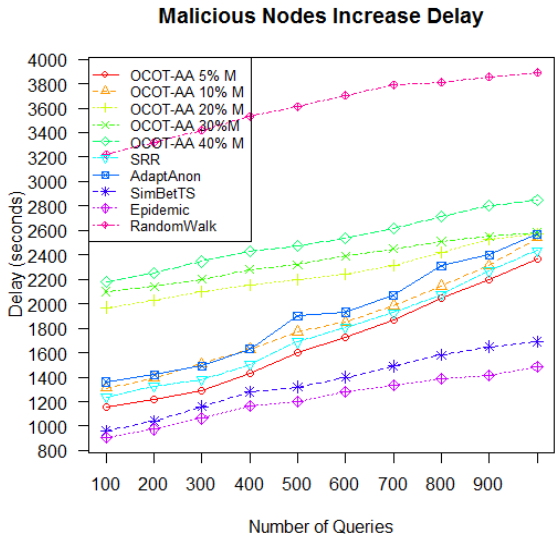

Figure 6. Delay vs. increasing number of queries

4.5.2. Length of the Obfuscation Path (L) and Diversity of Nodes on it (D)

Figure 7 shows that AdaptAnon [21] manages the highest K because of its multi-criteria selection of next hop anonymisers and lack of awareness of malicious nodes. It is followed by OCOT-AA 5%M and then OCOT-AA 10% M, OCOT-AA 20% M, OCOT-AA 30% M and OCOT-AA 40%M ranging from 5 to 3. The higher percentage of malicious nodes there are, the fewer nodes are able to test. This leads to shorter best effort K that is inversely proportional to the percentage of malicious nodes. SRR [19] manages the lowest K while SimBetTS [11] a higher than it but still lower than OCOT-AA 40% M. This is because SimBetTS has more opportunities to choose from when selecting the anonymisation nodes compared to SRR but still less opportunity than OCOT-AA 40% M whose efficient testing stage allows it to choose a diversity of next hops.

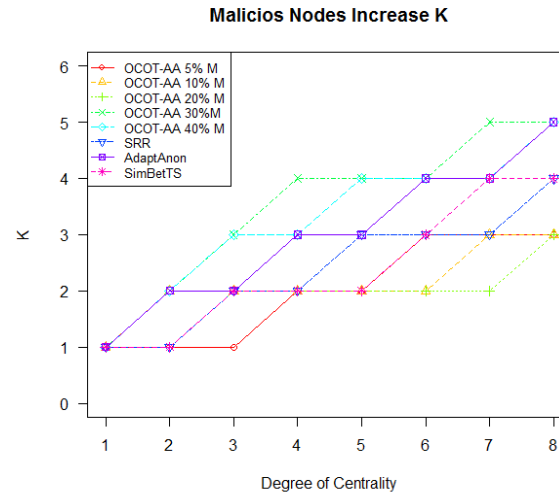

Figure 7. K vs increasing number of queries

Figure 7 shows the diversification factor of various percentages of malicious nodes over increasing number of queries. AdaptAnon manages the highest D because it specifically considers intelligent anonymisation that prevents it from over-utilising the same nodes for the anonymisation path. Next follow all OCOT-AA protocols as they manage to accurately and timely identify malicious nodes and are able to avoid them by using other nodes effectively.

We observe that all OCOT-AA protocols achieve very similar level of D that range between 65% and 75% for low sending rates and between 40% and 55% for high sending rates. AA 5%M has the highest level of diversification within them. Because SimBetTS and SRR use only contact and social role similarity respectively they manage low diversification of 61% and 52% for low sending rates and 21% and 15% for high sending rates. This is about two to five times less than what AdaptAnon achieves, and 40% to 30% lower than OCOT-AA with 40% malicious nodes.

4.5.3 Number of Tests Performed and Failed Anonymisation

We have analysed the total number of test performed over 30 days across 540 cabs for increasing number of malicious nodes. As the number of discovered "trusted"

nodes increases, increasing number of nodes have the opportunity to perform testing and the total number of tests grows. We observed that the average number of tests per hour, per car ranges from 38 to 130. This is reasonable frequency of testing in the face of unknown number and distribution of malicious nodes.

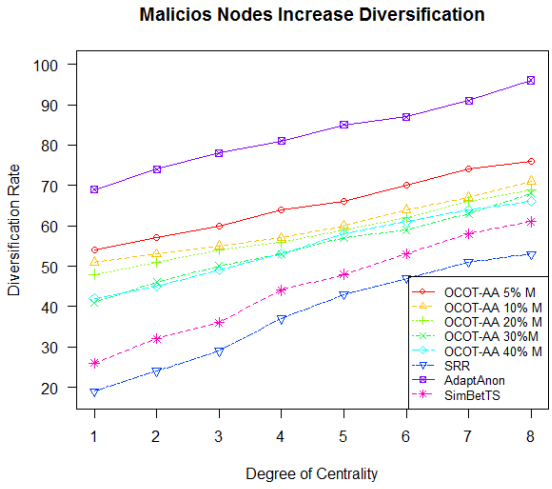

Figure 8. D rate vs increasing number of queries

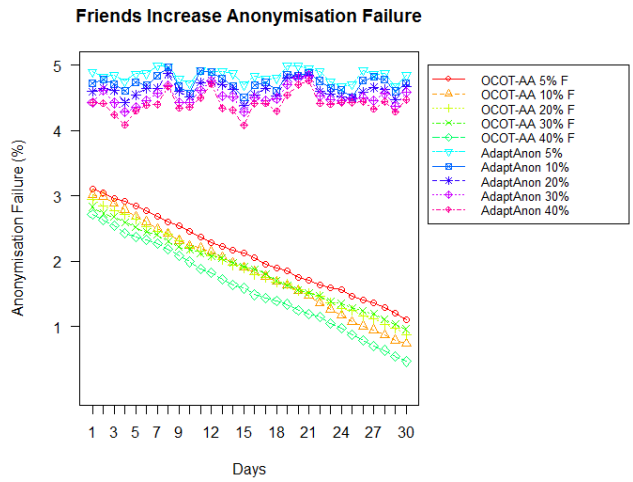

Figure 9. Anonymisation failure vs time

Figure 9 shows percentage of failed source rate of anonymisation for AdaptAnon [21] and OCOT-AA over time of 30 days in the presence of various percentages of "friend" nodes. For the purpose of clear analyses note that we fix the number of malicious nodes to 5% and that is the reason for the failure rates not being above 5%. We observe that for OCOT-AA there is no significant difference in the percentage of de-obfuscated nodes when the number of "friends" change but that there is a steep learning curve of the nodes that accurately detect malicious nodes and are able to avoid them as the time progresses (the failure rates drop from around 3% to 1%). For AdaptAnon (OCOT-AA without testing) we observe failed anonymisation rate that is proportional to the percentage of the malicious nodes we consider here. This is due to calculating failed anonymisation percentage as average per day across all nodes that are highly likely to meet all the malicious nodes.

Figure 10 shows the percentage of failed anonymisation for OCOT-AA and AdaptAnon [23] over the one month period and in the face of increasing proportion of malicious nodes. We define failed anonymisation as the real ID of the source being revealed to the destination. For OCOT-AA, we observe that scenario of OCOT-AA with 5% malicious nodes has around 3% of failure in the early stages of learning but quickly drops to 1% while OCOT-AA 40% malicious nodes scenario starts with failure rates of 5.2% but drops to 4%. This is because at all times OCOT-AA performs testing to detect and avoid malicious nodes. OCOT-AA 10%, 20% and 30% lie in between. This shows that our testing stage is efficient as the nodes reputation values are resolved quickly but also that the OCOT-AA is efficiently hiding the real source IDs for very different percentages of malicious nodes and at different stages of learning. The only reason for failure is when nodes with previous high reputation values turn malicious and fail to anonymise but their reputation status has not expired fast enough.

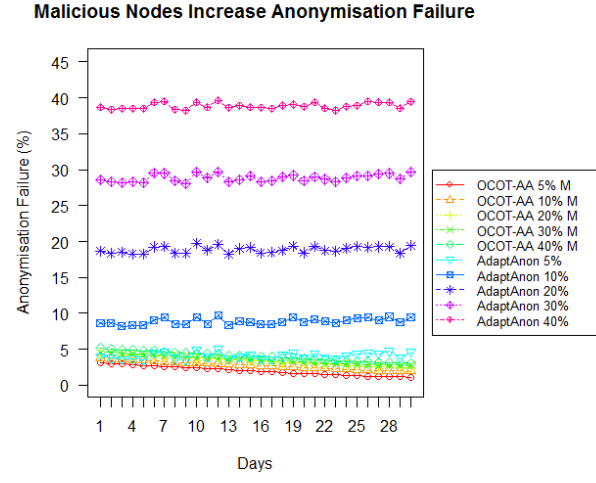

Figure 10. Anonymisation failure vs time

This is a reasonable trade-off between dramatically increasing testing cost versus decreasing 1% or 2% of anonymisation failure rates. Without testing, we show that AdaptAnon has much higher percentage of failed anonymisation ranging from 5% to 40% for increasing number of malicious nodes. This is due to AdaptAnon not being aware of malicious nodes and experiencing average anonymisation failure for all nodes proportional to the percentage of malicious nodes in the topology.

5. CONCLUSIONS

This paper proposed a novel reputation aware obfuscation framework (OCOT-AA) for mobile opportunistic networks that integrates to novel techniques: multi criteria fully distributed intelligent

anonymisation and opportunistic collaborative obfuscation testing. OCOT-AA allows successful source anonymisation even in the presence of large number of malicious nodes in the network. We assume that malicious nodes can intercept the messages from the senders while they are on their multi-hop path from the source to the destination and reveal sources' identity. Because, we assume that any node can become malicious at any time, our testing technique is fully localised and allows testing of any node by two nodes as long as they are trusted above a certain threshold. We use three types of implicit heuristics for intelligent anonymisation that allows responsiveness to the underlying topology, social interests and history of anonymisation behaviour. OCOT-AA manages high detection rate and avoidance of malicious nodes in the face of varying levels of malicious nodes and varying node degree centrality while it manages long anonymisation path, high diversification of anonymisation nodes, success ratios of delivered messages and low delays. We compare OCOT-AA to benchmark algorithms that exploit underlying connectivity analysis and social roles for building the "trusted obfuscation path" and analyse the impact of malicious node on four protocols with different decision making criteria for building the anonymisation path for sparse mobile environments. We show that OCOT-AA manages better trade-off and achieves better results in terms of ten criteria compared to other benchmark and competitive protocols.

REFERENCES

[1] Milena Radenkovic, Andrew Grundy, "Efficient and Adaptive Congestion Control for Heterogeneous Delay Tolerant Networks", Elsevier Ad Hoc Networks journal, April 2012

[2] Cab mobility traces, Exploratorium - the museum of science, the cabspotting project, http://cabspotting.org/

[3] E Bulut, B K Szymanski, Exploiting Friendship Relations for Efficient Routing in Mobile Social Networks, IEEE Transactions on Parallel and Distributed Systems, 2012

[4] E Bulut, B K. Szymanski: Friendship Based Routing in Delay Tolerant Mobile Social Networks. GLOBECOM 2010: 1-5

[5] A Tran, N Hopper, Y Kim. Hashing it out in public: common failure modes of DHT-based anonymity schemes. In Proc. OfWPES '09. ACM, NY, USA

[6] I. Parris, T. Henderson: Privacy-enhanced social-network routing. Computer Communications 35(1): 62-74 (2012)

[7] G. Bigwood, D. Rehunathan, M. Bateman, et al, CRAWDAD data set st_andrews/sassy (v. 2011-06-03)

[8] Mehul Motani, Vikram Srinivasan, Pavan S. Nuggehall, PeopleNet: engineering a wireless virtual social network, in Proceedings of ACM Mobicom 2005, 243-257, NY, USA

[9] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft. A socio-aware overlay for publish/ subscribe communication in delay tolerant networks. In ACM MSWiM'07:

[10] A-K PIETILÄINEN, C. DIOT, Dissemination in Opportunistic Social Networks: The Role of Temporal Communities, MobiHoc'12 2012

[11] E. Daly, M. Haahr, Social network analysis for information flow in disconnected Delay-Tolerant MANETs, IEEE Trans. Mob. Comp, 2009

[12] Asthana, S.; Kalofonos, D.N., "The Problem of Bluetooth Pollution and Accelerating Connectivity in Bluetooth Ad-Hoc Networks," Pervasive Computing and ommunications, PerCom 2005. Third IEEE International Conference on , vol., no., pp.200,207, 8-12 March 2005

[13] M. May, V. Lenders, G. Karlsson, and C. Wacha. Wireless opportunistic podcasting: implementation and design tradeoffs. In CHANTS'07: Proceedings of the second workshop on Challenged networks CHANTS, 2007

[14] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau. CRAWDAD data set cambridge/haggle (v. 2009-05-29). Downloaded from http://crawdad.cs.dartmouth.edu/cambridge/haggle, May 2009

[15] L. McNamara, C. Mascolo, and L. Capra. Media sharing based on colocation prediction in urban transport. In MobiCom'08: Proceedings, 2008

[16] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft. Opportunistic content distribution in an urban setting. In CHANTS'06: Proceedings of the 2006 SIGCOMM workshop on Challenged networks, 2006

[17] T. Hyyryl¨ainen, T. K¨arkk¨ainen, C. Luo, et al. Opportunistic email distribution and access in challenged heterogeneous environments. In CHANTS'07: Proceedings of the second ACM workshop on Challenged networks, 2007

[18] Greg Bigwood, Tristan Henderson: Bootstrapping opportunistic networks using social roles. WOWMOM 2011: 1-6

[19] S. Ioannidis, A. Chaintreau, and L. Massouli´e. Optimal and scalable distribution of content updates over a mobile social network. In Proceedings of IEEE INFOCOM, 2009

[20] Kanchana Thilakarathna, Aline Carneiro Viana, Aruna Seneviratne, Henrik Petander: Mobile social networking through friend-to-friend opportunistic content dissemination. MobiHoc 2013: 263-266

[21] M.Radenkovic, I. Vaghi, .S. Zakhary, A. Benslimane, "AdaptAnon: Adaptive Anonymity for Service Queries in Mobile Opportunistic Networks", in the Proceedings of IEEE ICC, Budapest, Hungary, June, 2013

[22] M Radenkovic, I Vaghi, Adaptive User Anonymity for Mobile Opportunistic Networks, ACM MobiCom CHANTS 2012.

[23] A. J. Khan, V. Subbaraju, A. Misra, S. Seshan, Mitigating the true cost of advertisement supported "free" mobile applications, HotMobile 2012

[24] A. Keränen, J. Ott, T. Kärkkäinen The ONE simulator for DTN protocol evaluation Proceedings of the 2nd International Conference on Simulation Tools and Techniques (SIMUTools), ICST, Brussels, Belgium (2009), pp. 1–10

[25] Lindgren, A. Doria, O. Schelén, Probabilistic routing in intermittently connected networks, in: Proceedings of the 1st International Workshop on Service Assurance with Partial and Intermittent Resources (SAPIR), 2004, pp. 239–254

[26] P. Hui, J. Crowcroft, E. Yoneki Bubble rap: social-based forwarding in delay tolerant networks, Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), ACM, New York, NY, USA (2008)

[27] Zhong Li, Cheng Wang, Siqian Yang, Changjun Jiang, Ivan Stojmenovic: Improving Data Forwarding in Mobile Social Networks with Infrastructure Support: A Space-Crossing Community Approach. CoRR abs/1307.7326 (2013)

[28] Afra Mashhadi, Pan Hui, "Proactive Caching for Hybrid Urban Mobile Networks",, Deutsche Telekom Laboratories, Germany, Paper ID: 1569334407

[29] K. Fall, S. Farrell: DTN: an architectural retrospective. IEEE Journal on Selected Areas in Communications 26(5) (2008)

[30] Anna Kaisa Pietiläinen, Christophe Diot: Dissemination in opportunistic social networks: the role of temporal communities. MobiHoc 2012: 165-174

[31] Greg Bigwood, Tristan Henderson: IRONMAN: Using Social Networks to Add Incentives and Reputation to Opportunistic Networks. SocialCom/PASSAT 2011: 65-72

[32] R15 Ólafur R. Helgason, Emre A. Yavuz, Sylvia T. Kouyoumdjieva, Ljubica Pajevic, and Gunnar Karlsson, "A Mobile Peer-to-Peer System for Opportunistic Content-Centric Networking", In Proc Mobiheld 2010, p 21-26

[33] Pan Hui, Jon Crowcroft, Eiko Yoneki: BUBBLE Rap: Social-Based Forwarding in Delay-Tolerant Networks. IEEE Trans. Mob. Comput. 10(11): 1576-1589 (2011

[34] S. Zakhary, M. Radenkovic, and A. Benslimane. Efficient Location Privacy-Aware Forwarding in Opportunistic Mobile Networks. *Accepted for publication in Transactions on Vehicular Technology, IEEE*, 2013

[35] Abdullatif Shikfa, "Security Issues in Opportunistic Networks" Mobiopp 2010 February 22-23, 2010

AUTHOR BIOGRAPHIES

Milena Radenkovic has received her PhD Degree from the University of Nottingham, UK and her Dipl Ing (Msc) from the University of Nis, Serbia. Her research spans the areas of mobile and delay tolerant networking, P2P systems, and their application to pervasive gaming, social networking, and environmental monitoring. She has been an Investigator of four EPSRC grants. Milena has organised and chaired over ten international IEEE and ACM conferences, served on many program committees and editorial boards and published in premium venues including Elsevier Ad Hoc Networks, IEEE Transactions on Vehicular Technology, ACM MC2R, IEEE ICC, IEEE WONS, IEEE Multimedia, MIT Press PRESENCE, ACM Multimedia, IEEE Multimedia, ACM VRST, ACM CCGRID. She has served as an Independent Expert for EU FP7 programme and has been Evaluator for both EU and EPSRC proposals.

Abderrahim Benslimane is a Technical International Expert at the French Ministry of Foreign and European affairs as Coordinator of the Faculty of Engineering at the French University of Egypt where he is also Head of the Informatics Research Center. He is Professor of Computer-Science at the Avignon University/France since 2001 from which he is now in secondment. He was attributed the French award of Scientific Excellency (2011-2014). He has been as Associate Professor at the University of Technology- Belfort-Montbéliard since 1994. He obtained the French title to supervise researches, 2000, Cergy-Pontoise University/France. He received the PhD degree, 1993, Franche-Comte University/France. His current research and teaching interests are in wireless and mobile networks. He has more than 130 refereed international publications.
He has several international collaborations for funded research projects and supervising PhD students. He is member of several editorial boards of international journals: IEEE Wireless Communication Magazine, Wiley WCMC, SCN, IJCS, Elsevier Ad Hoc, JNCA. He serves as General-Chair of the IEEE WiMob since 2008; he lunched and serves as General-Chair of iCOST and MoWNet since 2011. He serves as a Symposium co-chair/leader in many IEEE international conferences such as ICC, Globecom, AINA and VTC. He was GE of many special issues. He participates to the steering and the program committee of many IEEE international conferences. He is Board committee member, Vice-chair of Student activities of IEEE France section/Region 8, since 2008, Publication Vice-chair of the ComSoc TC of Communication and Information Security 2009-2011.

Derek McAuley has engaged in a career transitioning between academic research (Cambridge, Glasgow and Nottingham Universities), commercial research (founding member of labs for Microsoft, Marconi and Intel), and roles in several successful startups, most recently XenSource and Netronome. He is Professor of Digital Economy in the School of Computer Science at the University of Nottingham and Director of the Horizon Digital Economy Research Institute; his research has covered ubiquitous computing, computer architecture, networking, photonics, distributed systems and operating systems. He is a Fellow of the British Computer Society and member of the UKCRC, a computing research expert panel of the IET and BCS.